

# Extended Social Force Model-Based Mean Shift for Pedestrian Tracking Under Obstacle Avoidance

Xuguang Zhang<sup>1</sup>, Xufeng Zhang<sup>1</sup>, Yiming Wang<sup>2</sup>, Hui Yu<sup>2,\*</sup>

1. School of Electrical Engineering, Yanshan University, China E-mail: zhangxg@ysu.edu.cn

2. School of Creative Technologies, University of Portsmouth, Portsmouth, PO1 2DJ, U.K.

\*E-mail: hui.yu@port.ac.uk (corresponding author)

**Abstract:** It has been shown that the mean shift tracking algorithm can achieve excellent results in the task of pedestrian tracking. It empirically estimates the target position of the current frame by locating the maximum of a density function from the local neighborhood of the target position of the previous frame. However, this method only considers its past trajectory without taking into account the influence of the pedestrian's environment. In practice, pedestrians always keep a safe distance away from obstacles when programming their paths. To address the issue of obstacle avoidance, this paper proposes a novel extended social force model-based mean shift tracking algorithm, in which the pedestrian's environment is taken into full consideration. Firstly, in order to show how the environment impacts pedestrian movements from the viewpoint of force, an extended social force model is presented by considering the interaction between target and obstacle. Furthermore, according to characteristics of pedestrian tracking, directional weights and speed weights are introduced to adjust the strength of the force concerning the difference of individual perspectives and relative velocities. Finally, the initial target position is predicted by Newton's laws of motion and then the mean shift method is integrated to track the target position. Experiment results showed that this algorithm achieved an encouraging performance when an obstacle occurred. An object that moves fast or changes its moving directions quickly can also be robustly tracked in real time by using the proposed algorithm.

## 1 Introduction

Target tracking has been widely used in video surveillance systems [1, 2], intelligent human-computer interaction [3,7], robot vision navigation [4], intelligent transportation [5] and behavior analysis [6]. In recent years, an increasing number of researchers have been involved in the study of target tracking and a number of good target tracking algorithms have been proposed. Among them, the mean shift tracking algorithm attracts considerable attention due to its significantly low calculation consumption [32, 33]. In the pedestrian tracking process, the traditional mean shift tracking algorithm takes the target position of the current frame as the initial position of the next frame. This method is simple and straightforward but requires the candidate regions between two frames to be overlapping. In the situation that the pedestrian moves fast or changes directions quickly, or is under the circumstances of occlusion, the algorithm may be unstable or even fail to track.

To tackle this problem, many approaches have been proposed. In [8], the researcher used the method of least squares fitting to predict the pedestrian position based on its past trajectory. There are some other works that predict pedestrians based on instantaneous velocity. To some extent, these methods improve the algorithm's speed and accuracy. However, such approaches only consider the past trajectory

and ignore an important issue regarding human behavior. It is generally believed that people are driven by their future destination, as well as influenced by the environment; therefore, they adjust their trajectories at an early stage [9]. People always choose the most comfortable and convenient trajectories on the premise of keeping the safe and comfort psychological distance from the obstacles. That is to say, pedestrian environment is one of the dominant factors for its trajectory planning. Pedestrian simulation models usually quantify the pedestrian environment and use this factor to guide pedestrian movement. Thus, in order to consider the pedestrian environment in the algorithm, a social force model is introduced to the pedestrian tracking algorithm in this paper.

A number of pedestrian simulation models present detailed descriptions of pedestrian movements by focusing on an analysis of pedestrians' unique attributes. These models propose to solve the problem of pedestrian simulation by calculating the interaction of pedestrians. The most commonly used models include the social force model [10, 11], cellular automaton model [12, 13], discrete choice model [14, 15] and queuing network model [16]. Among them, Helbing's social force model is a simple but effective alternative, in which the influencing factors of individual motion are comprehensively modeled. Because of its rationality in modeling individual behavior, it has achieved great progress and attracted more and more attention in recent years. The social force model by Helbing [17, 18] is a computational model in which interactions between pedestrians are described by using the concept of a social force. It is based on the assumption that changes in human behavior can be explained in terms of social force; the pedestrian is treated as a particle driven by the force. A situation in which pedestrians intend to keep a safe distance from obstacles can be described as a function of the repulsive force imposed by the obstacle. Meanwhile, pedestrians are also attracted by their destinations so that they are expected to move round the obstacle to the target. It is this force that keeps pedestrians away from the obstacle at a comfortable psychological distance and avoids collision. In this paper, we propose a mean shift tracking algorithm for collision avoidance based on an extended social force model. Firstly, based on the theory of social force model, we propose an approach modeling the interaction between the pedestrian and obstacles by computing refined motion predictions and calculating the repulsive force exerted by the obstacles. In addition, knowing the fact that the influence on pedestrians imposed by the obstacles will be different due to the variations in individual perspective and relative velocity, an evolutionary social force model is presented, in which the directional weights and speed weights are considered. Finally, on the basis of calculating the repulsive force, the initial position of the target is predicted by Newton's laws of motion and then mean shift is utilized to track the target position. Thus, we can combine the traditional mean shift algorithm with the social force model to achieve a more reasonable tracking algorithm.

The paper is structured as follows: Section 2 introduces the theory of traditional mean shift tracking algorithm and social force model. Section 3 elaborates on the proposed approach for pedestrian tracking

under collision avoidance. Section 4 presents experimental results. Section 5 presents conclusions and future work.

## 2 Theoretical Analysis and Problem Formulation

### 2.1. Theory of mean shift tracking

The mean shift algorithm is a nonparametric density estimation method by which fast and optimal mode matching can be achieved [19, 20]. Let  $\hat{q}_u(y)$  be the color histogram of the target model; similarly, let  $\hat{p}_u(y)$  be the color histogram of the target candidate centered at  $y$ . Referring to [21], it can be defined as:

$$\hat{q}_u = C_h \sum_{i=1}^{n_h} k\left[\left\|\frac{y_0 - x_i}{h}\right\|^2\right] \delta[b(x_i - u)] \quad (1)$$

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k\left[\left\|\frac{y - x_i}{h}\right\|^2\right] \delta[b(x_i - u)] \quad (2)$$

where,  $k$  is the kernel function;  $h$  is the size of both kernel;  $C_h$  is the normalization factors;  $\delta$  is the Kronecker delta function;  $u$  represents the index of the color feature.

After the color histogram is obtained, the Bhattacharyya coefficient is employed as the similarity function, which is

$$\hat{\rho}(y) = \rho(\hat{p}(y), \hat{q}) = \sum_{u=1}^{m-1} \sqrt{\hat{p}_u(y) \cdot \hat{q}_u} \quad (3)$$

The bigger the coefficient, the larger the possibility that the target appears in the candidate area centered at  $y$ . Then the mean shift procedure is employed to find the maximum of density model in a local neighborhood. In this procedure, the kernel is recursively moved from the current location to the new location according to the relation

$$\hat{y}_{j+1} = \frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{\hat{y}_j - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{\hat{y}_j - x_i}{h}\right\|^2\right)} \quad (4)$$

where  $g(x) = -k'(x)$ , and the weight is computed as:

$$w_i = \sum_{u=0}^{m-1} \sqrt{\hat{q}_u / \hat{p}_u(y)} \delta[b(x_i) - u] \quad (5)$$

As for the initial position of this procedure, in the first frame a rectangular patch centered at  $y_0$  is manually initialized as the target model, and the iterations begin at  $y_0$ . For the rest of the frames, the target position of the previous frame is used as the initial position of the current frame [22]. This method requires the target regions in consecutive two frames having some overlap areas. If the target moves very

fast, the mean shift tracker's performance will be reduced dramatically. This work aims to overcome this limitation and takes the pedestrian environment into consideration. The details are presented in Section 3.2.

## 2.2. Traditional social force model

As in physics, social force is assumed to be the sum of several force terms which describes different factors that affect the behavior of the pedestrian simultaneously. The forces include psychological force, physical force and driving force. The corresponding expressions in [27, 28] are shown as follows:

$$\mathbf{f}_{ij}^{\text{psy}} = ak \cdot \exp[(r_{ij} - d_{ij}) / bk] \cdot \mathbf{n}_{ij} \quad (6)$$

$$\mathbf{f}_{ij}^{\text{phy}} = kg(r_{ij} - d_{ij}) \cdot \mathbf{n}_{ij} + Kg(r_{ij} - d_{ij}) \cdot \Delta v_{ji}^t \cdot \mathbf{t}_{ij} \quad (7)$$

$$\mathbf{f}_i^{\text{dri}}(t) = m_i \cdot \frac{v_i^0(t) \cdot \mathbf{e}_i^0(t) - \mathbf{v}_i(t)}{\tau_i} \quad (8)$$

The pedestrians are influenced by environmental changes during movement and adjust their state of motion. Pedestrians can gradually move to the destination with the premise of keeping a safe and comfort psychological distance from other objects. Psychological force is used by the model to express this phenomenon, as is described in equation (6). From this equation, pedestrians keep away from obstacles in order to reduce the risk of getting hurt. The smaller the distance, the more uncomfortable a pedestrian feels. This phenomenon can be described by a repulsive force  $\mathbf{f}_{ij}^{\text{psy}}$ . Pedestrians and obstacles are assumed to be in a circular shape with radii  $r_i$  and  $r_j$ . Then distance  $d_{ij}$  is given by the Euclidean distance between the centers, and  $r_{ij}$  is the sum of their radii. This force imposed by obstacles is a repulsive force, and the corresponding term  $\mathbf{n}_{ij}$  is the normalized vector pointing from  $p_j$  to  $p_i$ , which describes the direction of the force. In this equation,  $ak$  and  $bk$  are constants that determine the strength and range of the social interaction.

The  $\mathbf{f}_{ij}^{\text{phy}}$  in the equation (7) is physical force which only works in the case that contacts, extrusions or collisions between pedestrian and obstacle occur. And the  $k$  and  $K$  are constants, vector  $\mathbf{t}_{ij}$  is the unit tangential vector of the contacting circles,  $\Delta v_{ji}^t$  is the difference of the tangential velocities of the humans in contact, and it is derived by  $\Delta v_{ji}^t = (\mathbf{v}_j - \mathbf{v}_i) \cdot \mathbf{t}_{ij}$ . The function  $g$  is defined as:

$$g(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{else} \end{cases} \quad (9)$$

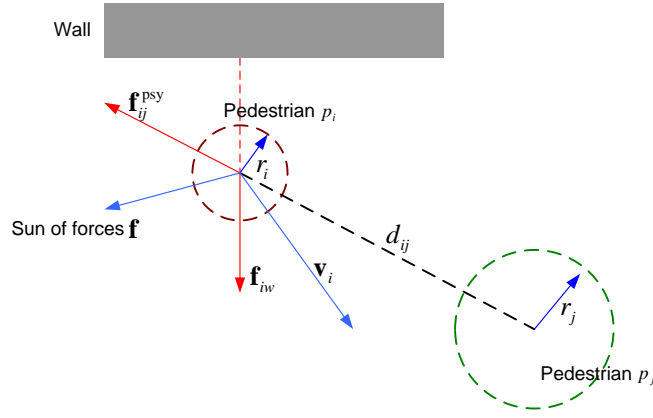
Yet, in real life, due to the anticipation and flexibility of pedestrian movements, the above situation does not usually occur. Therefore, we do not consider this force in our model.

The  $\mathbf{f}_i^{\text{dri}}$  in equation (8) is the driving force, which describes the relationship between pedestrian actual velocity and desired velocity. In the equation,  $m_i$  is the particle mass,  $v_i^0(t)$  is the magnitude of

desired velocity,  $\mathbf{e}_i^0(t)$  is the unit vector pointing to the desired target,  $\mathbf{v}_i(t)$  is the actual velocity and  $\tau_i$  is a time constant related to the relaxation time of the particle to achieve  $\mathbf{v}_i^0(t)$ . However, estimating desired velocity from the observation of a tracked person involves the problem of intention or activity recognition, which is an issue beyond the scope of this work. Since we focus on short-term human motion prediction, we make the weak assumption that the desired velocity of pedestrians is equal to their current velocity. While this assumption is simple, experiment results show that it is effective.

In the social force model, the dynamics of each pedestrian are assumed to be driven by these forces. In this way, the influence of the environment that affects pedestrian movements is calculated. The functions are plotted in Fig. 1. It is worth noting that, in this paper, we just use the social force model to improve the tracking performance when pedestrians encounter obstacles. Thus, the force exert by the wall is ignored in our method if the wall does not appear in front of the pedestrian as an obstacle.

The social force model represents the influence exerting to a pedestrian by an obstacle in the real scene. However, the position we gained by the target tracking method in an image is expressed as pixels. Therefore, the metrics in pixels should be converted to metrics in meters to calculate social force. That is to say, we should find how pixel values can be represented by distance in terms of meters in actual space. It is similar to spatial resolution. In this paper, we use a coefficient  $H$  to represent this relationship, i.e., one pixel in image space means  $H$  meters in actual space. The coefficient  $H$  can be calculated accurately by camera calibration. However, usually we should track a pedestrian in an uncontrolled camera. We do not know the parameter of the camera. In turn, in this paper we use a method to estimate the coefficient  $H$  approximately. For an image sequence, we choose one pedestrian, and calculate the shoulder width of this person. In order to set a shoulder width in this paper, we reference the elbow-to-elbow breadth of human data of 1960-1962 in the United States. In this reference, the definition of elbow-to-elbow breadth is the two elbows close to the body. The average breadth of men from 18-79 years old is 16.6 inches; the average breadth of women is 15.3 inches [39]. When a pedestrian is moving, the width of shoulder will be larger than elbow-to-elbow breadth because the pedestrian needs to shake the arm when walking. Therefore, we assume that the shoulder width is 50cm. In this paper, the visual angle that we captured the image sequences is from the top to bottom. In this visual angle the size of a target does not change dramatically. We thus think it is reasonable to use a fixed shoulder width here. Using this we can estimate the spatial resolution  $H$  approximately to calculate the velocity and social force of this pedestrian.



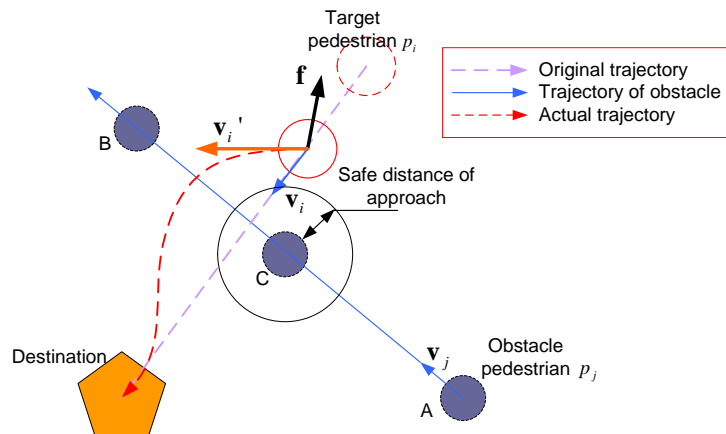
**Fig. 1** Forces of social force model

### 3 Methodology

The social force model is a pedestrian simulation model based on physical forces, in which interactions between pedestrians are calculated quantitatively in terms of social force. In this model, each pedestrian is treated as a particle, obeying kinematic equations and driven by the social force [25, 26]. In this paper, the social force model is integrated into mean shift and used for pedestrian tracking under the circumstance of collision.

#### 3.1. Pedestrian prediction under obstacle avoidance

In order to move towards a destination on the premise of avoiding collision, pedestrians will observe the surrounding environment ceaselessly, adjust the moving speed and direction whenever necessary and then move round the obstacle to the destination [23, 24]. In this paper, this process is simulated by assuming that a repulsive force between the target and obstacle exists. According to mechanical principles, this force will impose on the target and ensure the safe distance between target and obstacle (Fig. 2).



**Fig. 2** Mechanism of pedestrian movements under collision avoidance

When the target moves towards his destination, he will be influenced by the obstacle moving from position A to B. As the distance between them decreases, the effect of the force gradually becomes larger,

and it reaches the maximum values with the obstacle approach to position  $C$ . Driven by this force, the target will not move along his original path, but move round the obstacle to the destination. Thus, he can keep a safe distance when approaching the obstacle. The target will continue to move towards the destination after meeting the obstacle, and the force will reduce correspondingly. In this paper, an extended social force model is integrated to tackle the issue of pedestrian tracking under the circumstance of obstacle avoidance.

On the basis of calculating the social force, Newton's laws of motion are used to predict the initial position. The obstacle exerts a repulsive force to target, which is equivalent to constituting acceleration to the target. Trajectory of the target is predicted as follows:

$$y'_{j+1} = \hat{y}_j + \frac{v_n \cdot T + \frac{1}{2m_i} \cdot \|\mathbf{f}\| \cdot T^2}{H} \quad (10)$$

where  $\|\mathbf{f}\|$  is the norm of social force, with  $m_i$  being the pedestrian mass,  $\hat{y}_j$  being the target position of current frame,  $y'_{j+1}$  being the prediction position of the next frame,  $v_n$  being the current speed of the target,  $T$  being the time step and  $H$  being the spatial resolution.

### 3.2. Extended social force model

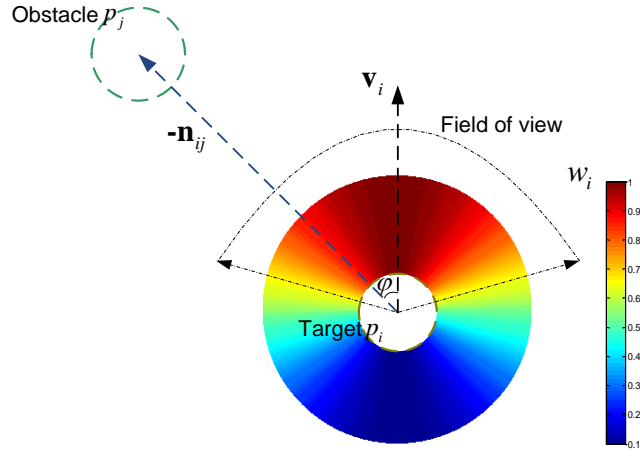
The traditional social force model is designed for a pedestrian simulation system. It has great potential for pedestrian tracking for real-world applications. In this section, an extended social force model is proposed.

#### 3.2.1 Directional weights

Given the limited field of view of humans, influences might not be isotropic. Generally speaking, pedestrians pay more attention to conditions in their immediate field of view, while the influence of conditions behind them is smaller than frontal conditions [29]. To change the strength and range of social force in terms of the different angle between desired direction and the direction of social force, the psychological force is formally expressed by scaling the forces with an anisotropic factor

$$w_i = \lambda + (1 - \lambda) \cdot \frac{1 + \cos \varphi}{2} \quad (11)$$

where  $w_i$  is the directional weights,  $\varphi$  is the angle between desired direction and opposite direction of social force;  $\lambda$  is a constant which ranges from 0 to 1 depending on the strength of the anisotropic factor.



**Fig. 3** Directional weights in improved social force model

(In the example, the constant  $\lambda$  in function 11 is 0.1, and red indicates good choices for velocity, blue signals “no-go”s.)

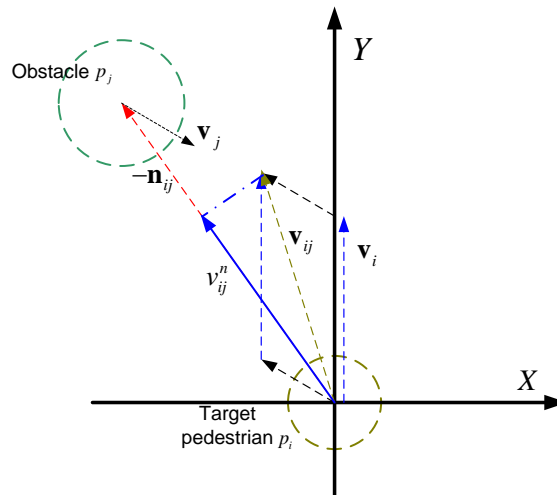
### 3.2.2 Speed weights

In real life, compared with objects approaching pedestrians slowly, responses of pedestrians will be earlier and more intense when objects approach pedestrians rapidly. Also, objects moving away from pedestrians will have minimal influence on pedestrians [30]. Let  $v_{ij}^n$  as the orthographic projection of relative velocity in direction  $-\mathbf{n}_{ij}$ , i.e.,  $v_{ij}^n = (\mathbf{v}_i - \mathbf{v}_j) \cdot -\mathbf{n}_{ij}$ . To make the model more compliable with the objective law of pedestrian movement, we make the strength and range of social force increase in direct proportion to  $v_{ij}^n$ , as shown in Fig. 4.

Equation (7) shows that the strength and range of social force are proportional to parameter  $bk$ . By setting  $ak=2000$ ,  $r_{ij}=0.4\text{m}$  and not considering the directional weights and speed weights, changes of social force with parameter  $bk$  are plotted in Fig. 5.

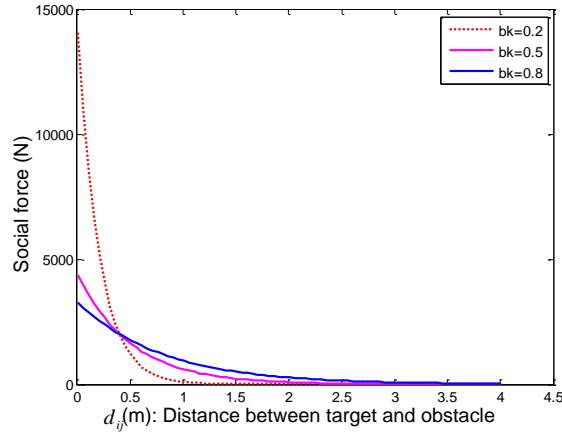
We want to improve the model and make social force increase in direct proportion to  $v_{ij}^n$ . The parameter  $bk$  is thus adjusted by the speed weights:

$$bk = ck + \delta v_{ij}^n \quad (12)$$



**Fig. 4** Speed weights in improved social force model





**Fig. 5** Changes of social force with parameter  $bk$

The parameter  $ck$  determines the weight when the relative velocity is zero. And  $\delta$  determines the weight of the relative speed, i.e., the impact of relative velocity on the force. The bigger it is, the greater the impact. When  $\delta$  becomes 0, we can assume that the social force model does not take the relative velocity into account.

The equation of the psychological force weighted by speed is described as follows:

$$\mathbf{f}_{ij}^{\text{psy}} = ak \cdot \exp[(r_{ij} - d_{ij}) / (ck + \delta v_{ij}^n)] \cdot w_i \cdot \mathbf{n}_{ij} \quad (13)$$

### 3.3. Extended social force model-based mean shift tracking

In this paper, the social force model was integrated into mean shift tracking algorithm. Let the mean shift algorithm begin iterations at position  $\mathbf{y}'_{j+1}$  calculated by equation (10); that is,  $\mathbf{y}'_{j+1}$  is the initial position of  $j+1$  frame. Combining equation (4) and equation (10), the improved mean shift tracking algorithm is:

$$\hat{\mathbf{y}}_{j+1} = \frac{\sum_{i=1}^{n_h} x_i w_i g(\|\frac{\mathbf{y}'_{j+1} - \mathbf{x}_i}{h}\|^2)}{\sum_{i=1}^{n_h} w_i g(\|\frac{\mathbf{y}'_{j+1} - \mathbf{x}_i}{h}\|^2)}, j = 0, 1, 2, \dots, n \quad (14)$$

The tracking process of this algorithm is as follows:

Algorithm 1 Improved mean shift tracking
<p><b>Input:</b> Image sequence</p> <ol style="list-style-type: none"> <li>1: Read the first frame, mark the target and the obstacle</li> <li>2: Calculate the color histogram of target model <math>\hat{q}_u</math> and initialize target position centered at <math>\mathbf{y}_0</math> according to equation (1, 2)</li> <li>3: Run traditional Mean Shift algorithm until it converges to the final location of target <math>\hat{\mathbf{y}}</math>.</li> <li>4: <b>Repeat</b></li> <li>5: Read the next frame and predict target position <math>\mathbf{y}'</math> based on extended social force model</li> <li>6: Use predicted position <math>\mathbf{y}'</math> as initial position.</li> <li>7: Run improved mean shift tracking algorithm according to equation (14).</li> <li>8: <b>Until</b></li> </ol> <p>It reaches the final frame.</p>

## 4 Experimental Results

To test performance and demonstrate the versatility of the approach, a series of experiments was conducted on both a public access database and our own database. Four sets of experiments were designed to test tracking performance when the pedestrians avoided dynamic obstacles and static obstacles. All experiments were performed on the MATLAB platform. The positions of obstacles in each experiment were obtained by using a traditional mean shift tracking algorithm in advance. For the sequences captured by ourselves, each frame of the video sequence was in BMP format with a resolution of  $640 \times 480$  pixels. As for the public ETH Walking Pedestrians (EWAP) dataset [34] [35], the resolution of each image is  $720 \times 576$  pixels. In this paper, we focus on solving the problem where the target changes its trajectory drastically when obstacles occur. In some application fields of target tracking, the target moves very fast, or the frame rate of the video is low, whereas the traditional mean shift method cannot track the target well. Thus, this problem has been considered in the proposed method. In our experiments, some sequences were captured in a low frame rate. Some of the sequences used in our experiments although short were sufficient for testing the proposed method. The target moves from one point to another through obstacle avoidance in the sequence, but it does not affect the performance of the proposed method.

Some researchers used a prediction algorithm based on the mean shift and least square method, which can handle the occlusion effectively and can successfully track the fast moving target. We compare the tracking performance between the traditional mean shift tracking algorithm (MS), mean shift algorithm combined with the least squares method (LMS), the mean shift algorithm combined with Kalman filtering (KMS) and the algorithm proposed in this paper (FMS). In MS, the target is represented by probability density functions of color feature with a size of  $16 \times 16 \times 16$  for convenience. In LMS, the initial position is obtained by learning a linear least square regressor based on the position of the previous eight frames. In KMS, the initial position of the target is predicted by the Kalman filter. Referred to [18, 31], parameters for social forces are determined empirically in this paper (Table 1). If  $ak$  and  $ck$  are of big values, the repulsive force will become too large, which may lead to failure in tracking. Otherwise, if the parameters are of small values, the impact of force will become too weak. In this case, it can be treated as the model in [9] where only the velocity is considered.

**Table 1** Parameters for social forces

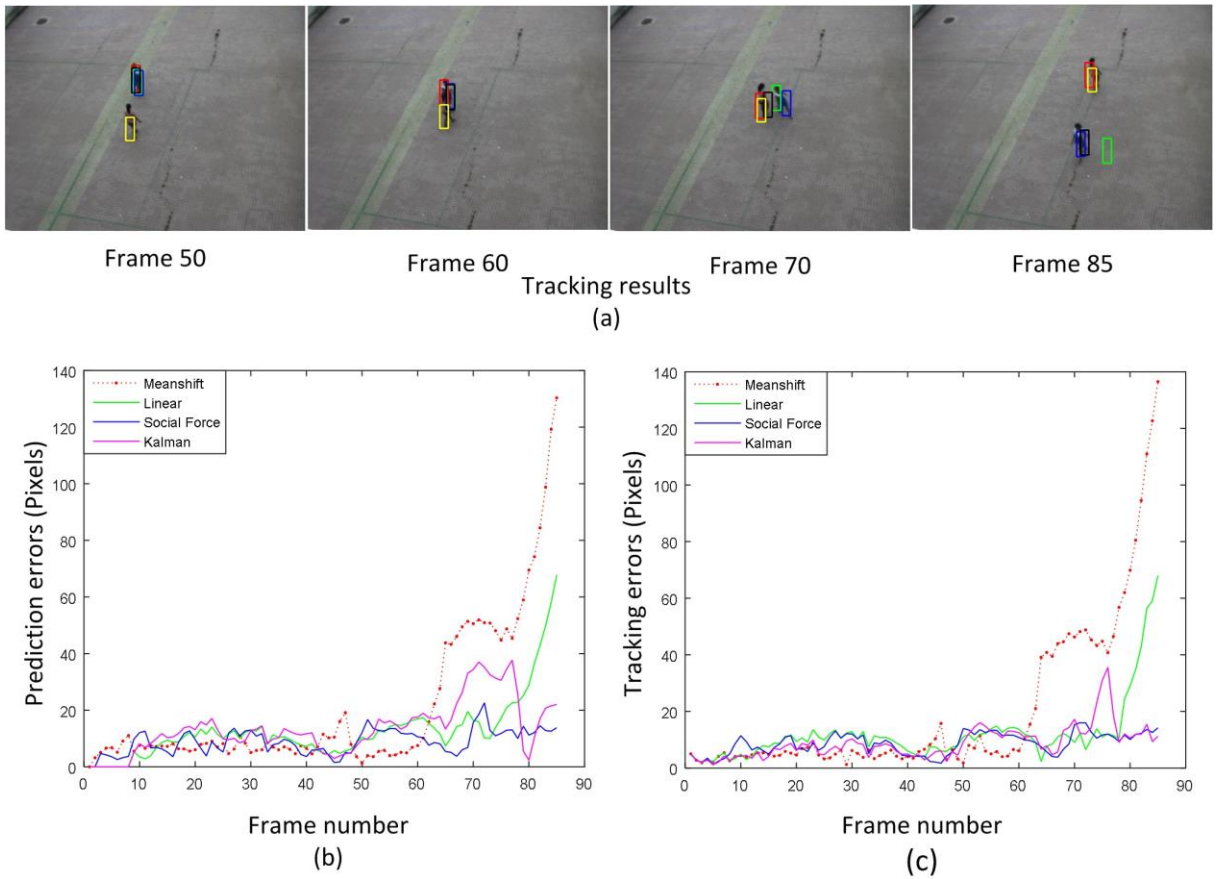
Parameters	$ak$	$ck$	$\delta$	$\lambda$	$m$
1st experiment	2000	0.5	0.0175	0.1	50kg
2st experiment	2000	0.5	0.0175	0.1	50kg
3rd experiment	2000	0.25	0.0375	0.1	50kg
4th experiment	2000	0.5	0.0175	0.1	50kg

#### 4.1. Experiments for pedestrian tracking under dynamic collision avoidance

The first two experiments are designed to test the tracking performance of our approach when the pedestrians avoid dynamic obstacles. To test whether the algorithm can handle the occlusion effectively, such a scenario is designed considering that two pedestrians move alone and their directions of motion are orthographic.

##### 4.1.1 Experiment one under dynamic collision avoidance

In the first experiment, the pedestrian is treated as the obstacle moves from southwest to northeast. At the same time, the pedestrian is treated as the target moves from northwest to southeast. When the pedestrian is close to the obstacle, that person will weave around it. Alternatively, occlusions occur when they meet each other in the center of the space. The dataset with 15 FPS consists of 85 frames. In the simulation, rectangles in red, green, black and blue show the target position obtained by using algorithms MS, LMS, KMS and FMS, respectively, and the yellow rectangle shows the position of the obstacle. Frames with the occlusion are showed in Fig. 6(a).



**Fig. 6** Comparison of experiment one

The ground truth of target trajectory is obtained by manual annotation. In order to compare the prediction accuracy of each algorithm, prediction errors (i.e., Euclidean distances of initial positions from their true positions) are plotted in Fig. 6(b). Similarly, tracking errors (i.e., Euclidean distances of tracking positions from their true position) are plotted in Fig. 6(c) to illuminate the comparison of tracking

accuracy. In addition, mean square errors of tracking results for a whole sequence [36] can be calculated by equation (15):

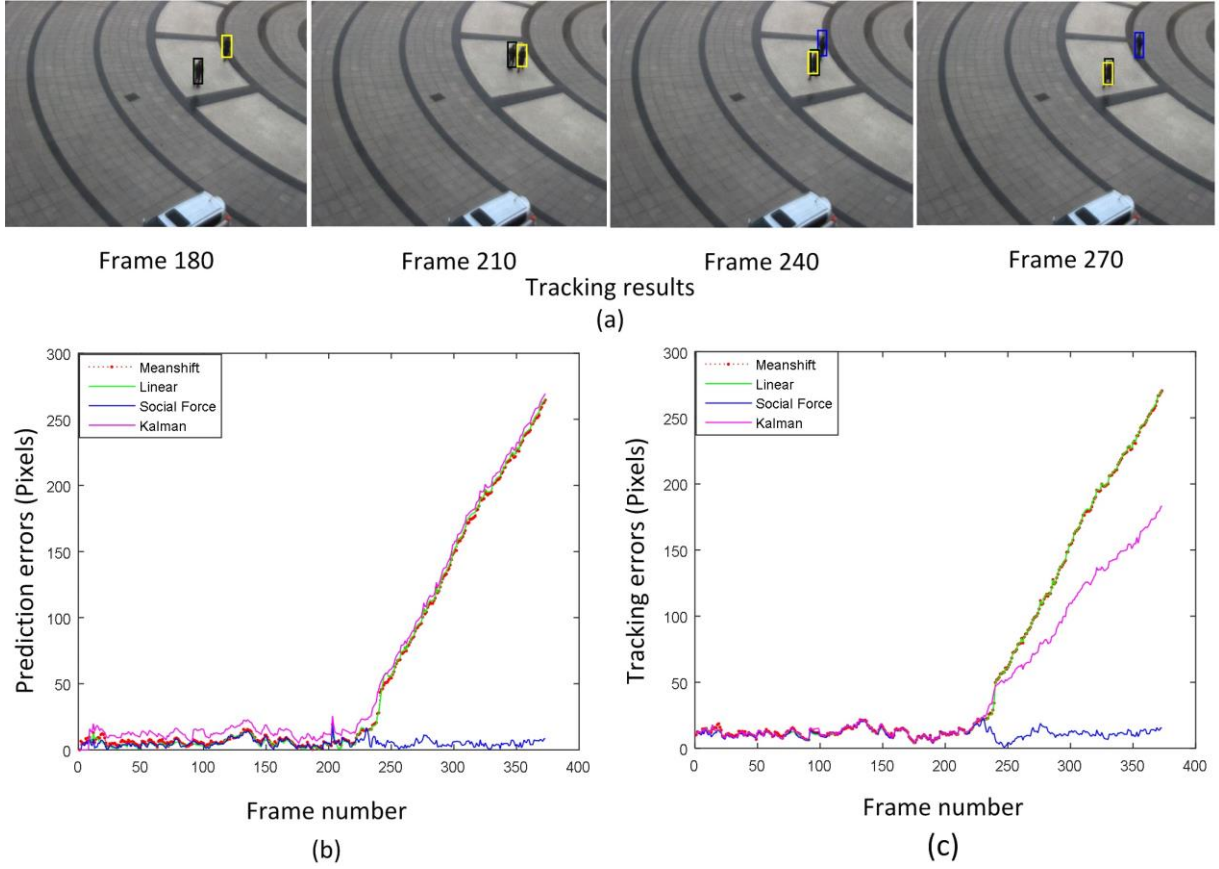
$$m = \sqrt{\frac{\sum_{i=1}^k [(r_i - t_i)^2]}{k}} \quad (15)$$

where  $m$  is the mean square of tracking errors,  $k$  is the frame number of this dataset,  $r_i$  is the tracking result of frame  $i$ , and the  $t_i$  is the ground truth of this frame.

In this sequence, for MS, LMS, KMS and FMS algorithms, their mean square error are 34.92, 16.48, 10.59 and 9.47, respectively. It is obvious that the FMS algorithm has a better tracking performance. It can be seen that before the pedestrians meet each other, prediction errors of all three algorithms are less than 20 pixels. In this period, the tracking error is low enough to ensure successful tracking. However, after the 60th frame, the target will be close enough to the obstacle. Due to occlusions and quick changes in moving directions, the prediction error of algorithm MS is larger than 30 pixels, which makes it impossible to find the right position near the initial location. Then, it fails in tracking the target after frame 70. As for the LMS algorithm combined with least square fitting, its prediction errors nearby the meeting position (i.e., frame 60) are slightly larger than the algorithm in this paper. However, as is noticed in Fig. 6(b), its tracking errors increase dramatically in the final phase of tracking, which indicates that linear fitting fails in predicting the target when the target changes its moving directions. For the KMS algorithm, the prediction position of the target is still based on the past trajectory and velocity of pedestrian. Therefore, when the target changes its moving directions, the prediction errors and tracking errors of the KMS algorithm are larger than FMS. The FMS benefits from the social force model that fully considers the pedestrian environment. As a result, the predictions errors maintain at less than 20 pixels on the whole sequence. This method outperforms the others at the final tracking stage, which suggests its robust tracking property.

#### 4.1.2 Experiment two under dynamic collision avoidance

In order to further test the performance of the algorithm under dynamic collision avoidance, we collect another sequence with 373 frames. In this sequence, the pedestrian moves from southwest to northeast, and the target moves from northwest to southeast. Frames with the occlusion are shown in Fig. 7(a), and markings are same as that in Fig. 6(a).

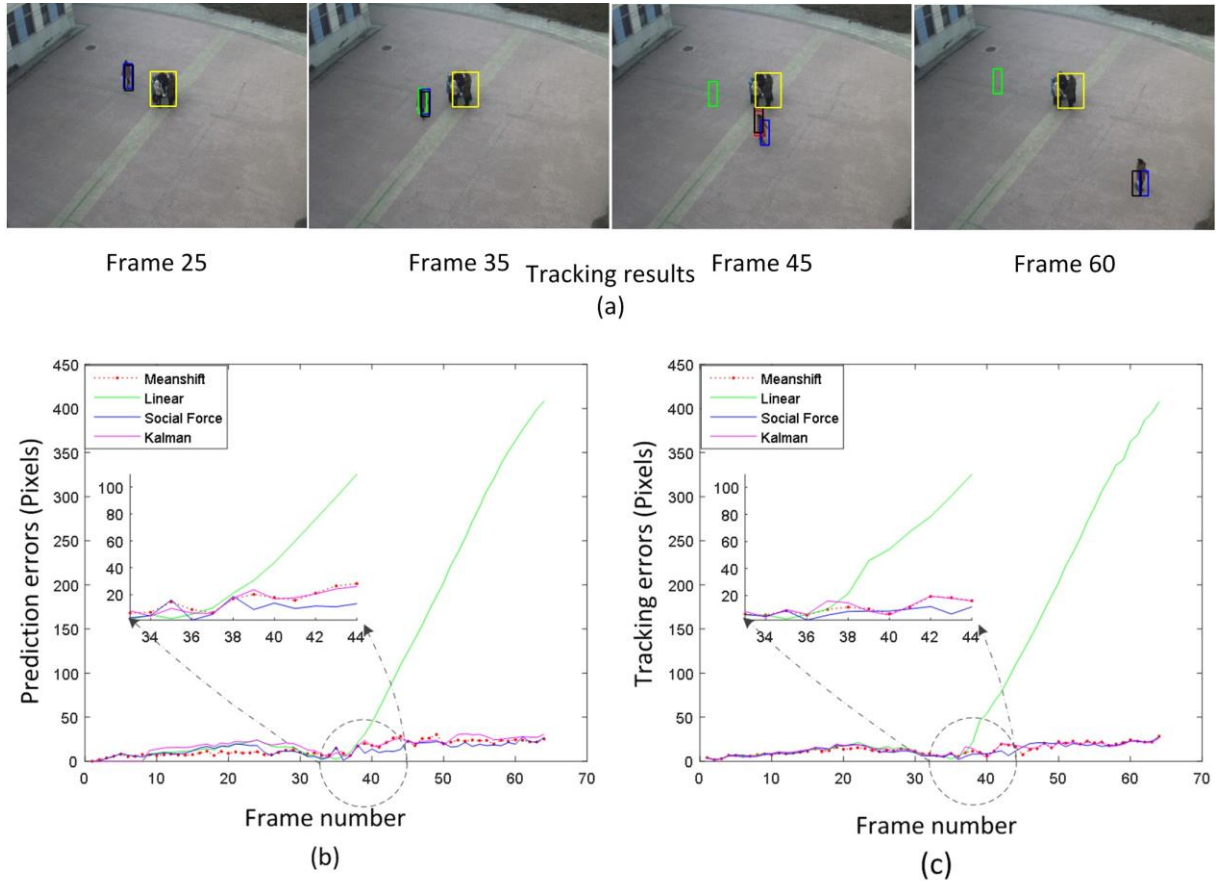


**Fig. 7** Comparison of experiment two

In this experiment, the prediction errors and tracking errors of the four methods are plotted in Fig. 7(b) and Fig. 7(c). As in the first experiment, the performance of the proposed FMS method is better than the others. Before the pedestrians meet each other, prediction errors of all three algorithms are small, so every method can track target correctly. However, after the pedestrians meet each other, only the FMS algorithm can track the target successfully. The reason that the other three methods failed is because the appearance features between object and obstacle were similar. Therefore, the position information of the object is important. Since the proposed FMS algorithm fully considers the interaction between pedestrians, predictions errors maintain about 15 pixels in the whole sequence. What is more, for MS, LMS, KMS and FMS algorithms, the mean square errors are 104.5, 104.7, 72.5 and 11.8, respectively.

#### 4.2. Experiment for pedestrian tracking under static collision avoidance

Considering the convenience of image sequences capturing, in this experiment a static pedestrian group is assumed as obstacles, such as street lamp and telephone booth. The target moves from northwest to southeast and occlusions occur when he is close to the obstacle located in the midpoint of the path. The data set with 7.5 FPS consists of 64 frames. Frames near the occlusion are shown in Fig. 8(a), and markings are the same as Fig. 6(a).



**Fig. 8** Comparison of experiment three

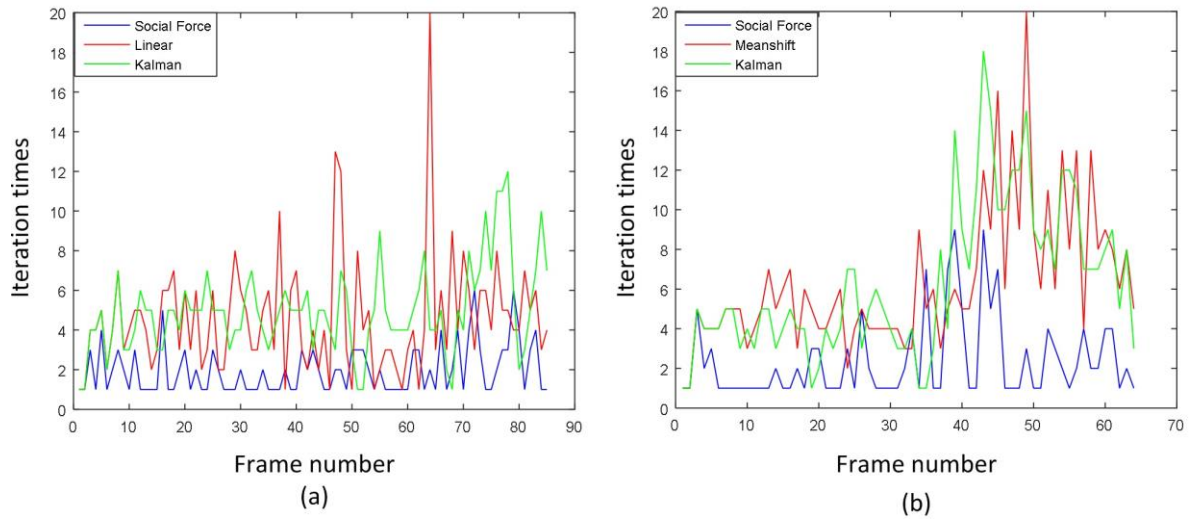
Similar to experiment one, in this experiment the prediction errors and tracking errors of the four methods are plotted in Fig. 8(b) and Fig. 8(c). Before the 30th frame, the target moves in a straight line and all the algorithms can track the target successfully. It can be seen from the figure that the prediction errors are low enough at these frames, so mean shift can continue tracking until it finds the true position successfully in the vicinity of the initial position. Thereafter, since linear fitting is unable to predict the curvilinear path if the target weaves form around the obstacle, algorithm LMS loses the target quickly at frame 38. Its mean square of tracking error is 163.21. The MS and KMS method also can track target successfully in this sequence. However, the initial positions gained by these two methods are still further away from the true position compared with algorithm FMS, which may lead to bigger tracking errors. Then, for MS, KMS and FMS, the mean square of tracking error is 14.44, 14.91 and 13.91, respectively.

Combined with the result of the first experiment, in all experiments, only the algorithm FMS tracks the target successfully with a relatively low error. These results show that the proposed algorithm is more stable and accurate.

#### 4.3. Comparison of iteration times

The speed of the algorithm is an important evaluation criterion on the premise of tracking successfully. In order to compare the speed of algorithms that can track successfully, the iteration times of all frames are plotted in Fig. 9.





**Fig. 9** Comparison of iteration times

The iteration time for the first sequence is shown in Fig.9(a), and Fig.9(b) shows iteration times for the third sequence. For the second sequence, the iteration times are not calculated here because MS, LMS and KMS algorithm cannot track the target.

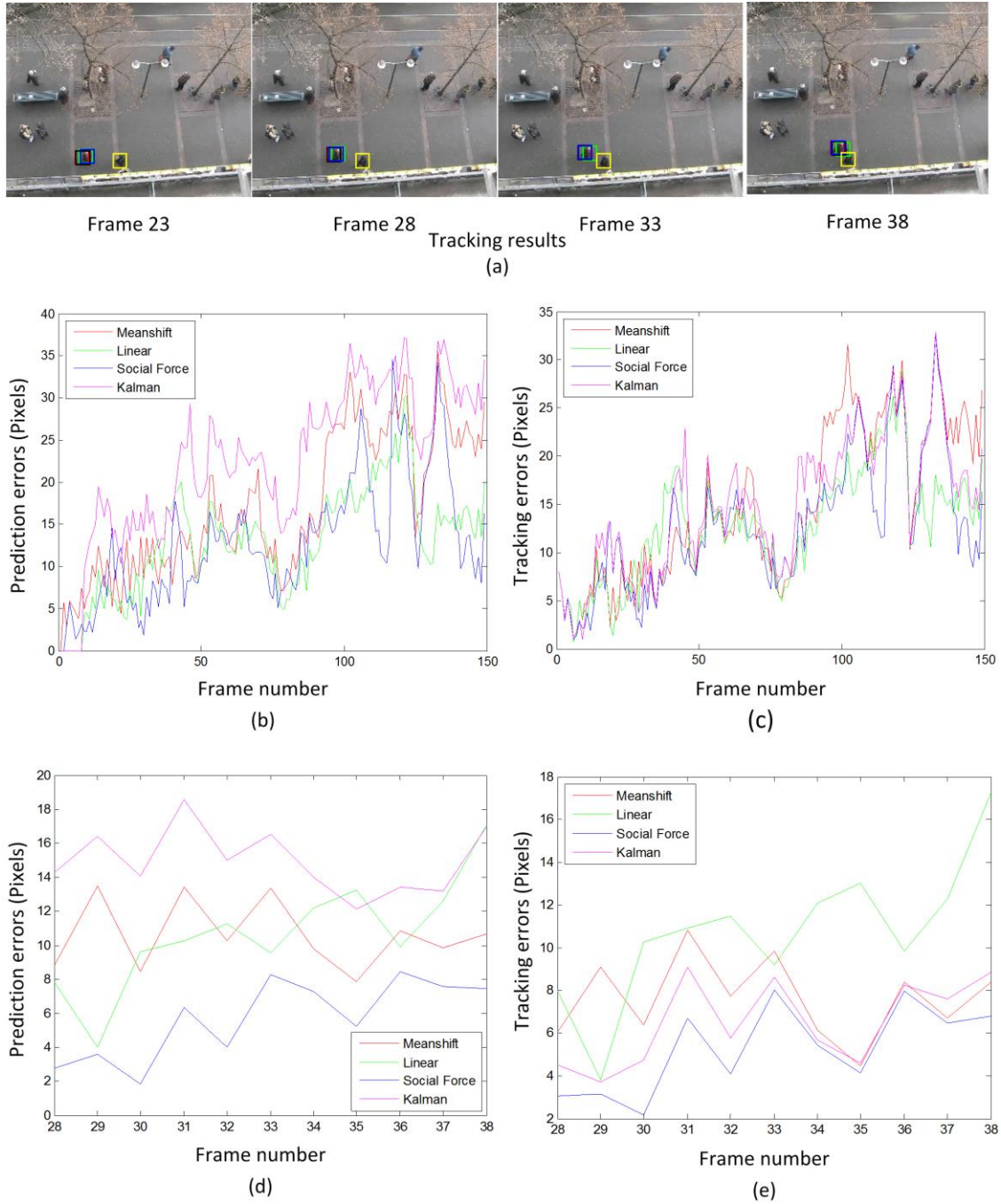
It is shown that initial positions of algorithm FMS are closer to the true target position compared with the others, which is conducive to improve tracking speed. The advantage in speed is even more obvious at the frames under occlusion. In the first experiment, after the 60th frame, the iteration time of algorithm FMS is 3 or 4, while LMS and KMS are more than 6 or even 14 times. Using the algorithm LMS, KMS and FMS, the average iteration times of the whole first sequence are 3.04, 4.96 and 1.95, respectively. Compared with the former, the iterative times of the latter reduce by 35.9% and 60.7%. In the second experiment, the improvement is greater. After the 45th frame, the number of iteration times of algorithm FMS is 3 and the other algorithms are more than 10. As for the entire second sequence, their average numbers are 6.33, 6.28 and 2.36, each. The algorithm in this paper improves the speed by 62.7% and 62.4%.

What is worth noticing is that, nearby the 15th frame of the first experiment and the 40th frame of the second experiment, the number of iteration times of the proposed algorithm is bigger than the others. This is because pedestrian movements will be influenced by the pedestrian's individual psychological status, which may change at any time. Therefore, the prediction of the target position will be fluctuation.

#### 4.4. Experimental results using public dataset

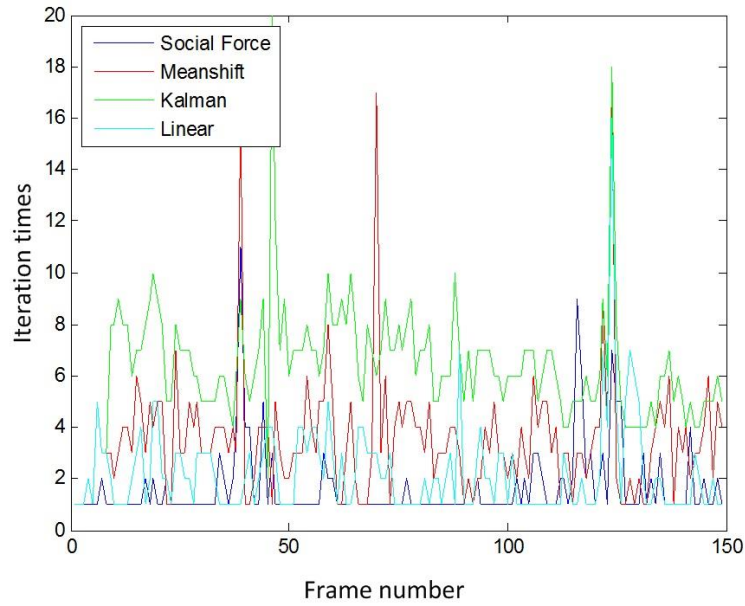
In order to demonstrate the performance of the proposed method further, we added one experiment using a sequence gained from the ETH Walking Pedestrians (EWAP) dataset. Since the algorithm is proposed for tracking a single pedestrian in obstacle avoidance situation, we extracted a short sequence from the long video. The video we used is called “seq\_hotel” in the EWAP dataset. We extracted 150 images from frame 7282 to frame 7431 to form a short sequence. The frame rate is 25FPS. In order to

show the experimental results clearly, we renumbered the sequence from frame 1 to frame 150 (i.e., we set the frame 7282 as frame 1). In this sequence, the obstacle pedestrian moves from right to left at the bottom of the image, and the target pedestrian walks from left to right. The obstacle avoidance behavior occurs from frame 25 to frame 40, approximately. The parameters used in this experiment are the same as experiment 1 (Table 1). Fig. 10(a) shows the tracking results of different methods. The tracking results are marked as the same of Fig. 6(a). The prediction errors and tracking errors are shown in Fig. 10 (b) and (c). In Fig. 10(d) and (e), we highlight prediction errors and tracking errors during the two pedestrians meeting. Fig. 11 shows the comparison of iteration times of the four methods.



**Fig. 10** Comparison of the experiment of hotel sequence





**Fig. 11** Comparison of the iteration times of hotel sequence

In the seq\_hotel sequence, the collision between the two pedestrians is not severe. In turn, all four methods can track the target during the whole sequence. However, in the stage of pedestrian avoidance, the tracking errors and the prediction errors of the proposed algorithm are lower compared with the other three methods. The proposed FMS method gains the least mean square error in the whole sequence compared with MS, LMS and KMS methods. The mean square errors of the whole sequence of MS, LMS, KMS and FMS methods are 17, 14, 16.5 and 13.5, respectively. During the obstacle avoidance phase, the mean square errors are 6.9, 9.8, 6.4 and 4.5. In addition, tracking performance reduces in the later stage of the sequence for all of the four methods, because we have not taken into account the scale change in the mean shift tracker. We also compare the average number of iterations of the four methods to evaluate the efficiency of the proposed method. The average numbers of iterations of the four methods are 3.5, 2.2, 6.4 and 1.6. The proposed method can track the target faster than the other three methods.

## 5 Conclusions and Future Work

In order to make the mean shift algorithm more effective and reasonable in the task of tracking pedestrians in an unconstrained environment, a novel approach that combines the mean shift with the social force model is proposed in this work for obstacle avoidance. Firstly, based on the traditional social force model, directional weights and speed weights are employed to model the changes of perspectives and relative velocity. Then, the influence of environment on the pedestrian movements is quantified by using the extended social force. Finally, the initial position is predicted by Newton's laws of motion and then mean shift is utilized to track the target. Experiment results show that this algorithm has superior performance on pedestrian tracking accuracy under static and dynamic collision avoidance compared to LMS, KMS and traditional mean shift algorithm, which claims a better property that the proposed method will track the target that moves fast or changes its directions quickly. As for computing speed, the

proposed method can track a target faster than the other three methods. Consequently, our method outperforms the other methods in both tracking accuracy and computational cost.

Future work will focus on improving the optimization and adaptation of parameters. It is challenging when considering more features in the tracking task. Furthermore, employing the pedestrian tracking method to resolve some questions about crowd behavior analysis and transportation management is also a potentially valuable work [37] [38].

## Acknowledgments

This research was supported by National Natural Science Foundation of China (No. 61271409) China Postdoctoral Science Foundation (No. 2012M510768, No. 2013T60264), Natural Science Foundation of Hebei Province, China (No. F2013203364), and China Scholarship Council (No. 2011813018).

## References

- [1] Torabi, A., Massé, G., Bilodeau, G.A.: ‘An iterative integrated framework for thermal–visible image registration, sensor fusion, and people tracking for video surveillance applications’, *Computer Vision and Image Understanding*, 2012, **116**, (2), pp. 210–221
- [2] Rao, Y.B.: ‘Automatic vehicle recognition in multiple cameras for video surveillance’, *The Visual Computer*, 2015, **31**, (3), pp. 271–280
- [3] Mekonnen, A.A., Leraslea, F., Herbulota, A.: ‘Cooperative passers-by tracking with a mobile robot and external cameras’, *Computer Vision and Image Understanding*, 2013, **117**, (10), pp. 1229–1244
- [4] Lee, D., Kim, G., Kim, D., Myung, H., Choi, H.T.: ‘Vision-based object detection and tracking for autonomous navigation of underwater robots’, *Ocean Engineering*, 2012, **48**, pp. 59–68
- [5] Pai, C.J., Tyan, H.R., Liang, Y.M.: ‘Pedestrian detection and tracking at crossroads’, *Pattern Recognition* 2004, **37**, (5), 1025–1034
- [6] Yan, X., Kakadiaris, I.A., Shah, S.K.: ‘Modeling local behavior for predicting social interactions towards human tracking’, *Pattern Recognition*, 2014, **47**, (4), pp. 1626–1641
- [7] Zhou, W, Fei, Z., HU, H. S., Liu, L., Li, J., Smith, P., Real-time Object Subspace Searching Based on Discrete Searching Paths and Local Energy. *Inter. Journal of Automation and Computing*, vol. 13, no. 2, 99-107, 2016.
- [8] Zheng, Y.Z., Wang, H.Y., Guo, Q.X.: ‘A Novel Mean Shift Algorithm Combined with Least Square Approach and Its Application in Target Tracking’, In: *Proc. IEEE Int. Conf. on Signal Proc.*, 2012, pp. 1102–1105
- [9] Pellegrini, S., Ess, A., Schindler, K.: ‘You’ll Never Walk Alone: Modeling Social Behavior for Multi-target Tracking’, In: *Proc. IEEE Int. Conf. on Computer Vision*, 2009, pp. 261–268
- [10] Zeng, W.L., Nakamura, H., Chen, P.: ‘A Modified Social Force Model for Pedestrian Behavior Simulation at Signalized Crosswalks’, *Procedia - Social and Behavioral Sciences*, 2014, **138**, pp. 521–530
- [11] Saboia, P., Goldenstein, S.: ‘Crowd simulation: applying mobile grids to the social force model’, *The Visual Computer*, 2012, **28**, (10), pp. 1039–1048

- [12] Jian, X.X., Wong, S.C., Zhang, P., Choi, K., Li, H., Zhang, X.N.: 'Perceived cost potential field cellular automata model with an aggregated force for pedestrian dynamics', *Transportation Research Part C: Emerging Technologies*, 2014, **42**, pp. 200–210
- [13] Henein, C.M., White, T.: 'Microscopic information processing and communication in crowd dynamics', *Physica A: Statistical Mechanics and its Applications*, 2010, **389**, (21), pp. 4636–4653
- [14] Kuwano, M., Zhang, J.Y., Fujiwara, A.: 'Dynamic discrete choice model for multiple social interactions', *Transportation Research Board*, 2011, **2231**, pp. 68–75
- [15] Robin, T., Antonini, G., Bierlaire, M., Cruz, J.: 'Specification, estimation and validation of a pedestrian walking behavior model', *Transportation Research Part B: Methodological*, 2009, **43**, (1), pp. 36–56
- [16] Roy, D., Krishnamurthy, A., Heragu, S., Malmberg, C.: 'Queuing models to analyze dwell-point and cross-aisle location in autonomous vehicle-based warehouse systems', *European Journal of Operational Research*, 2015, **242**, (1), pp. 72–87
- [17] Helbing, D., Farkas, I., Vicsek, T.: 'Simulating Dynamical Features of Escape Panic', *Nature*, 2000, **407**, (6803), pp. 487–490
- [18] Helbing, D., Buzna, L., Johansson, A., Werner, T.: 'Self-organized Pedestrian crowd Dynamics: Experiments, Simulations, and Design Solutions', *Transportation Science*, 2005, **39**, (1), pp. 1–24
- [19] Bousetouane, F., Dib, L., Snoussi, H.: 'Improved mean shift integrating texture and color features for robust real time object tracking', *The Visual Computer*, 2013, **29**, (3), pp. 155–170
- [20] Siyit, R., Ding, Y.G.: 'Research and realization of Mean Shift Tracking Algorithm', *Journal of Computational Information Systems*, 2014, **10**, (7), pp. 3057–3064
- [21] Maalouf, A., Larabi, M.C., Nicholson, D.: 'Offline quality monitoring for legal evidence images in video-surveillance applications', *Multimedia Tools and Applications*, **73**, (1), pp. 189–218
- [22] Li, Y., Liang, S., Bai, B.D., Feng, D.: 'Detecting and tracking dim small targets in infrared image sequences under complex backgrounds', *Multimedia Tools and Applications*, 2014, **71**, (3), pp. 1179–1199
- [23] Löhner, R.: 'On the modeling of pedestrian motion', *Applied Mathematical Modeling*, 2010, **34**, (2), pp. 366–382
- [24] Kim, S., Guy, S.J., Hillesland, K., Zafar, B., Gutub, A.A., Manocha, D.: 'Velocity-based modeling of physical interactions in dense crowds', *The Visual Computer*, 2014, doi: 10.1007/s00371-014-0946-1
- [25] Mehran, R., Oyama, A., Shah, M.: 'Abnormal crowd behavior detection using social force model', In: *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2009, pp. 935–942
- [26] Curtis, S., Zafar, B., Gutub, A., Manocha, D.: 'Right of way', *The Visual Computer*, 2013, **29**, (12), pp. 1277–1292
- [27] Johansson, F., Peterson, A., Tapani, A.: 'Waiting pedestrians in the social force model', *Physica A: Statistical Mechanics and its Applications*, 2015, **419**, pp. 95–107
- [28] Seer, S., Rudloff, C., Matyus, T., Brändle, N.: 'Validating Social Force based Models with Comprehensive Real World Motion Data', *Transportation Research Procedia*, 2014, **2**, pp. 724–732
- [29] Johansson, A., Helbing, D., Shukla, P.K.: 'Specification of a Microscopic Pedestrian Model by Evolutionary Adjustment to Video Tracking Data', *Advances in Complex Systems*, 2008, **10**, (4), pp. 271–288
- [30] Gao, Y., Luh, P.B., Zhang, H., Chen, T.: 'A modified social force model considering relative velocity of pedestrians', In: *Conf. IEEE Int. Conf. on Automation Science and Engineering*, 2013, pp. 747–751

- [31] Parisi, D.R., Gilman, M., Moldovan, H.: 'A modification of the Social Force Model can reproduce experimental data of pedestrian flows in normal conditions', *Physica A: Statistical Mechanics and its Applications*, 2009, **338**, (17), pp. 3600–3608
- [32] Beyan, C., Temizel, A.: 'Adaptive mean-shift for automated multi object tracking', *IET Computer Vision*, 2012, doi: 10.1049/iet-cvi.2011.0054
- [33] An, X., Kim, J., Han, Y.: 'Optimal colour-based mean shift algorithm for tracking objects', *IET Computer Vision*, 2012, doi: 10.1049/iet-cvi.2013.0004
- [34] S. Pellegrini, A. Ess, K. Schindler, L. van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," 2009 IEEE 12th International Conference on Computer Vision, pp. 261–268, Sept. 2009.
- [35] "Ethz - computer vision lab: Datasets." <http://www.vision.ee.ethz.ch/datasets/index.en.html>.
- [36] Yin, F., Dimitrios, M., Sergio, A. V.: 'Performance evaluation of object tracking algorithms', *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, 2007
- [37] Kaysi, I., Alshalalfah, B., Shalaby, A., Sayegh, A., Sayour, M., Gutub, A. "Users' Evaluation of Rail Systems in Mass Events: Case Study in Mecca, Saudi Arabia", *Transportation Research Record: Journal of the Transportation Research Board*, 2013, Issue 2350, pp. 111-118.
- [38] Abdelgawad, H., Shalaby, A., Abdulhai, B., Gutub, A., Microscopic modeling of large-scale pedestrian–vehicle conflicts in the city of Madinah, Saudi Arabia, *Journal of Advanced Transportation*, 2014, 48(6), pp.507–525.
- [39] Stoudt HW, Damon A, McFarland R, et al. Weight, height, and selected body dimensions of adults. Vital and health statistics series 11. DHEW Publication (PHS) 8. Rockville, MD: National Center for Health Statistics, 1969